

ПРОЕКТ OPEN UKRAINIAN CITATION INDEX (OUCI): ІДЕЯ, ПРИНЦИП РОБОТИ ТА ПЕРСПЕКТИВИ РОЗВИТКУ

Сергій НАЗАРОВЕЦЬ,

заступник директора з наукової роботи
Державної науково-технічної бібліотеки України,
м. Київ

Open Ukrainian Citation Index (OUCI) – проект створення пошукової системи і бази даних наукових цитувань, які надходять від усіх видань, що використовують сервіс Cited-by від Crossref та підтримують Initiative for Open Citations, який розробляється фахівцями Державної науково-технічної бібліотеки України за дорученням Колегії Міністерства освіти і науки України.

Вигоди від реалізації проекту OUCI: новий онлайн-інструмент спростить пошук наукової інформації в Інтернеті, пришвидшить процес відбору видань для публікації авторам, полегшить пошук рецензентів редакторам наукових журналів, мотивуватиме українських видавців наукової продукції дотримуватися міжнародних рекомендацій, покращить представлення українських видань у світовому інформаційному просторі, відкриє наукометристам додаткове джерело даних для проведення досліджень, підвищить інформаційну культуру користувачів, зокрема стимулюватиме надійно ідентифікувати джерела у списках використаної літератури (вказувати DOI), що спростить процес перевірки джерел і допоможе у боротьбі з плагіатом у академічному середовищі та з іншими формами порушень прав інтелектуальної власності.

Попри те, що ідея та принципи роботи OUCI є доволі простими, бібліотекарям ДНТБ України часто доводиться спростовувати хибні твердження окремих представників української наукової спільноти щодо різноманітних аспектів роботи майбутньої пошукової системи і бази даних цитувань. У цій статті я коротко поясню ідею створення і функціонування OUCI, а також спростую деякі помилкові твердження та побоювання потенційних майбутніх користувачів OUCI.

Аналіз цитувань між науковими публікаціями важливий, адже він дозволяє вченим дізнатися, де, ті чи інші, наукові результати були використані, де вони були підтверджені або ж заперечені, де дослідження отримали подальший розвиток. Процес встановлення цитатних зв'язків технічно складний, адже вчені використовують різноманітні стилі оформ-

млення бібліографічних списків, які потрібно не тільки точно представляти у машиночитаному форматі, крім того, алгоритми повинні вміти коректно розпізнавати різні елементи бібліографічного опису. Сьогодні найточнішими інструментами для аналізу наукових цитувань вважають комерційні продукти Web of Science Core Collection та Scopus, за доступ до яких наукові установи (які можуть собі це дозволити) щорічно сплачують десятки тисяч доларів. Проте далеко не всі учасники наукової спільноти переконані, що такий стан повинен залишатися незмінним.

Переважна більшість видавців наукової літератури передає метадані своїх видань до агенції Crossref для реєстрації DOI (Digital Object Identifier) – унікальних цифрових ідентифікаторів об'єктів, що спрощують ідентифікацію публікацій та убезпечують від втрати посилань при зміні веб-адрес ресурсів. Також Crossref заохочує видавців скористатися безкоштовним сервісом Cited-by (<https://www.crossref.org/services/cited-by/>), який дозволяє відстежувати цитування публікації: члени Crossref, які беруть участь у Cited-by, надають інформацію про документи, що процитовані у публікаціях, для яких вони реєструють DOI, і, водночас, можуть отримати інформацію про те, які документи цитують їхні видання. Реєстраційна агенція Crossref надає безперешкодний доступ до бази метаданих, однак лише за умови, що видавці погодяться відкрити власні метадані. Тому дедалі більшого поширення у науковому світі набуває Initiative for Open Citations (I4OC) – співпраця між науковими видавництвами, дослідниками та іншими зацікавленими сторонами з метою сприяння необмежених доступності даних про цитування в науковій літературі (<https://i4oc.org/>). Якщо до запуску ініціативи був відкритий приблизно лише 1% даних наукових цитувань, то зараз I4OC підтримало багато провідних наукових видавництв (American Geophysical Union,

Association for Computing Machinery, BMJ, Cambridge University Press, Cold Spring Harbor Laboratory Press, EMBO Press, Royal Society of Chemistry, SAGE Publishing, Springer Nature, Taylor & Francis, Wiley), і, станом на березень 2019 р., вже відкрито 54% наукових цитувань (43,2 мільйони записів DOI з відкритими списками використаної літератури).

Будь-який видавець, що використовує сервіс Cited-by від Crossref, може долучитися до Initiative for Open Citations, просто звернувшись електронним листом до Crossref з проханням відкрити його метадані. У Порядку формування Переліку наукових фахових видань України передбачено, що одним із критеріїв включення видання до категорії Б переліку є «присвоєння кожному опублікованому матеріалу міжнародного цифрового ідентифікатора DOI», а участь у безкоштовній програмі Cited-by від Crossref також заохочується Міністерством освіти та науки України. Саме на використанні цих нових ініціатив базується принцип роботи Open Ukrainian Citation Index (OUCI): наукові видавці отримують DOI від Crossref для своїх публікацій, використовують сервіс Cited-by та підтримують Initiative for Open Citations, що дозволить отримати відкриту базу даних наукових цитувань. Подальша реалізація проекту OUCI передбачає написання скрипта та інтерфейсу, за допомогою якого користувачі зможуть зручно проводити пошук наукових публікацій у відкритій базі наукових цитувань.

Давайте розглянемо і спростуємо низку помилкових тверджень про OUCI, які бібліотекарі ДНТБ України виявили в Мережі за перші місяці роботи над проектом.

Open Ukrainian Citation Index – це новий наукометричний показник.

Ні. Open Ukrainian Citation Index (OUCI) – це пошукова система і база даних наукових цитувань, які надходять від усіх видань, що використовують сервіс Cited-by від Crossref та підтримують

Initiative for Open Citations.

OUCI — це локальна база українських журналів, яка відстежуватиме лише цитування українських журналів. Це аналог локальних іноземних реферативних баз, таких, наприклад, як РИНЦ.

Ні. Розуміючи важливість функціонування національних реферативних баз даних наукової літератури, команда розробників OUCI, водночас, усвідомлює, що національні бази даних цитувань дають надто неповне відображення наукової комунікації. Тому ми відразу відмовилися від такого підходу, і пошукова система і база даних наукових цитувань OUCI міститиме метадані усіх видань, що використовують сервіс Cited-by від Crossref та підтримують Initiative for Open Citations.

OUCI — це інструмент оцінки якості наукових журналів, що уможливорює різноманітні «накрутки».

Ні. OUCI покликаний спростити пошук наукових публікацій, привернути увагу редакцій до проблеми повноти та якості метаданих українських наукових видань, дозволить бібліометристам вільно вивчати зв'язки між авторами та документами з різних наукових дисциплін, зокрема, в галузях суспільних та гуманітарних наук.

OUCI — це інструмент для приховування плагіату та поширення псевдонауки.

Ні. Відповідно до принципу роботи OUCI, в ідеалі, в майбутньому кожна позиція списку використаної літератури повинна буде містити DOI, відтак пошукова система OUCI навпаки полегшує процес перевірки джерел, які були використані в процесі написання роботи.

Для чого створювати OUCI, якщо Google Scholar індексує більшість наукових документів та цитувань?

На відміну від комерційних баз даних, Google Scholar значно краще охоплює наукову літературу різними мовами, проте вміст бази даних Google Scholar непрозорий, система індексує нерцензовані документи та цитування, а її комп'ютерний алгоритм незахище-

ний від маніпуляцій.

База OUCI містить метадані наукових видань, що отримують DOI від Crossref та враховує цитування між публікаціями (DOI to DOI), тож знайдені неточності, або метадані видань, що публікують рукописи без належного рецензування, можна буде видалити з результатів пошуку (наприклад, вказавши префікс DOI видання).

Показники OUCI можуть використати для хибних оцінок та маніпуляцій.

В Україні та світі є багато прикладів некоректного використання метрик, проте, проблема не в самих базах даних чи інструментах, а в людях, які не розуміють і хибно інтерпретують наукометричні показники. Такі ініціативи, як Сан-Франциська декларація про оцінку наукових досліджень, Лейденський маніфест, Metric Tide закликають до відповідального застосування наукометричних показників, і ми також рекомендуємо дотримуватись цих принципів.

Переважає більшість україномовних наукових журналів — поганої якості. Для чого їх індексувати?

Свобода слова і вільне вираження своїх поглядів і переконань гарантуються Конституцією України. Для розвитку вітчизняного наукового потенціалу та для створення простору якісної публічної комунікації вчених застосовується Перелік наукових фахових видань України.

Розробники OUCI переконані, що саме рецензування (розгляд наукових публікацій вченими, які є фахівцями у тій галузі науки, що й оцінювана ними публікація) є найкращим механізмом контролю якості наукових досліджень.

Ми вважаємо, що пропозиція не індексувати українські наукові видання — це такий ж нонсенс, як забороняти українським авторам представляти препринти на ArXiv.org. Якщо ж Вам відомі видання, які не проводять належного рецензування рукописів, але видають свої публікації за наукові — зверніться до головного редактора, засновника, видав-

ця, членів редколегії, колег-науковців у Вашій галузі з вимогою припинити підтримувати це видання.

Укладачі комерційних баз даних, таких як Scopus та Web of Science Core Collection, проводять відбір наукових видань. Чому в OUCI не планується запроваджувати критерії для відбору видань?

Ми приєднуємося до розбудови нової справедливої інфраструктури, де кожен, хто зацікавлений у поширенні наукових знань, зможе представити спільноті своє видання, розвинути навички експертної оцінки і скористатися доступом до наукових цитувань, щоб вивчати зв'язки між документами.

Якщо ви уважно переглянете списки використаної літератури публікацій, які представлені у базах Scopus чи Web of Science Core Collection, то зауважите, що у них згадуються багато наукових видань, що не включені у Scopus та/або Web of Science Core Collection, однак, які були використані у процесі написання робіт, що повинні були пройти добросовісне рецензування. Відповідно, сам лише факт представлення, або непередставлення видання у певних реферативних базах даних, не може слугувати критерієм науковості видання.

Проте для зручності користувачів OUCI, які звикли покладатися на експертні висновки інших ресурсів, ми плануємо запропонувати в інтерфейсі фільтри, які дозволять проводити пошук

документів лише серед видань, що представлені в Scopus, Web of Science, PubMed, DOAJ тощо (з дотриманням авторських прав).

Таким чином, пошукова система і база даних наукових цитувань OUCI повинна спростити вченим і бібліотекарям пошук та аналіз наукових документів. Також активна підтримка Initiative for Open Citations покращить представлення українських наукових видань у спеціалізованих пошукових системах (таких, наприклад, як Dimensions, Lens.org, Ifindr, Scilit), що може розширити їх читацьку аудиторію.

Написання скрипта OUCI — складне технічне завдання, однак, для цілковитого успіху проекту потрібно також дбати про повноту та якість метаданих наукових видань (вказувати DOI у списках літератури до статей, ORCID авторів, реферати публікацій, назви організацій, що фінансували дослідження тощо). Тому, якщо Ваш науковий журнал досі ще не підтримував Initiative for Open Citations, скористайтеся нашим відкритим практичним посібником (<https://doi.org/10.5281/zenodo.2554825>), щоб дізнатися, як скористатися сервісом Cited-by, як відкрити свої метадані, як передати у Crossref списки використаної літератури з допомогою форми Simple Text Query, або через Metadata Manager BETA, та долучайтеся до розвитку відкритої системи наукової комунікації.

А